深度学习综述

李海新

（北京化工大学信息科学与技术学院 北京 100020）

摘要：人工智能的核心课题之一是神经网络与深度学习，它们模仿人脑的工作原理，通过多层次的神经元连接来从数据中挖掘有价值的知识和规律。神经网络的研究始于上世纪四十年代，经过几次起伏和革新，目前已经涵盖了多个种类和领域，如卷积神经网络、循环神经网络、语音识别、计算机视觉和自然语言处理等。深度学习是指利用多层神经网络来处理复杂的非线性问题，它依赖于海量的数据和计算资源，以及高效的训练和优化技术。深度学习在近几年取得了令人惊叹的进展，但也存在着一些难题和挑战，如模型解释性、泛化能力、安全性和可靠性等。深度学习仍是一个充满活力和前景的研究领域，有望为人类的智能和生活开辟更多的机会和可能。本文将从神经网络到深度学习来简要介绍一下部分类型的神经网络结构以及部分深度学习的模型结构。

关键词：神经网络 深度学习 卷积神经网络 自编码器

分类号： TP393

Deep Learning Survey

Haixin Li

(School of Information Science and Technology, Beijing University of Chemical Technology,Beijing 10020,China)

Abstact：One of the core topics of artificial intelligence is neural networks and deep learning, which imitate the working principle of the human brain and use multi-level neural connections to mine valuable knowledge and rules from data. The research of neural networks started in the 1940s and went through several ups and downs and innovations. It now covers many types and fields, such as convolutional neural networks, recurrent neural networks, speech recognition, computer vision and natural language processing. Deep learning refers to using multi-layer neural networks to solve complex nonlinear problems. It relies on massive data and computing resources, as well as efficient training and optimization techniques. Deep learning has achieved amazing progress in recent years, but also faces some difficulties and challenges, such as model interpretability, generalization ability, security and reliability. Deep learning is still a vibrant and promising research field, which is expected to open up more opportunities and possibilities for human intelligence and life. This article will briefly introduce some types of neural network structures and some deep learning model structures.

Keywords: Neural networks Deep learning Convolutional neural networks AE

1 引言

神经网络的概念最早出现在1943年，当时两位学者McCulloch和Pitts基于神经元的结构和功能，建立了一个数学模型，称为MP模型。这个模型为后续的神经网络研究奠定了基础。在20世纪50年代末到60年代初，Rosenblatt在MP模型的基础上，增加了一种学习机制，设计了一个单层的感知器模型，这是神经网络的第一个实际应用。然而，单层感知器模型存在一个局限性，就是无法处理线性不可分的问题。直到1986年，Rumelhart等人提出了一种多层的前馈网络，使用误差反向传播的算法进行训练，这种网络叫做BP网络，它能够克服单层感知器模型的缺陷，解决一些更复杂的问题[1]。

神经网络是模仿人脑的神经元的工作原理，对输入的信息进行处理和学习。神经网络能够通过学习数据，建立起复杂的非线性的模型，从而实现对数据的分类和判别。

神经网络的典型结构包括输入层、隐藏层和输出层，隐藏层的个数可以根据需要调整。每一层的节点都通过权重连接，表示节点之间的相关性，每个节点都有一个激活函数，用来输出映射后的值。利用这些网络模型，可以完成函数逼近、优化求解、模式识别、目标分类等任务，所以，神经网络在工业生产、人工智能、机器人等领域的信息处理中有着广泛的应用[1]。

神经网络有许多种模型，比如前馈神经网络(FFNN)、径向基神经网络(RBF)、卷积神经网络(CNN)、自编码器(AE)、DNN、极限学习机(ELM)等。Nguyen Van Cuong等在2021年提出了一种基于DNN(Deep neural network)的轴承故障诊断方法来提升特征提取的性能，从而得到了更好的故障诊断性能[2]。Shahabodin Afrasiabi等提出了一种基于单模块Gabor滤波器的卷积神经网络(GCNN)，相较于SVM、LVQ、FFNN等来说拥有更好的性能[3]。

2 神经网络

神经元是生物神经系统的生物基础，大量的神经元相互连接组成了神经网络，神经网络具有感知和学习能力。为了建模需要，将神经元的功能数学化，建立神经元的数学模型，在神经元上作用非线性激活函数组成非线性网络系统，整体上具有模仿动物神经网络的行为特征。

2.1 神经元

神经元是构成神经网络的最基本的单元，它模拟了生物中大脑的神经元结构以及神经元之间互相建立联系，最后完成思考的特性。经典的神经网络结构一般包含输入层、中间层（或隐藏层）和输出层，隐藏层可根据实际情况设置多个。两层间的节点以权重相连，代表两节点间的轻重关系，每个神经元节点通过一种特定的激励函数映射后输出。运用这些网络模型可实现函数拟合、优化处理、模式识别、目标分类等功能，因此，神经网络广泛应用于工业生产过程、人工智能、机器人等领域的信息处理中。神经元的结构一般如图1所示。

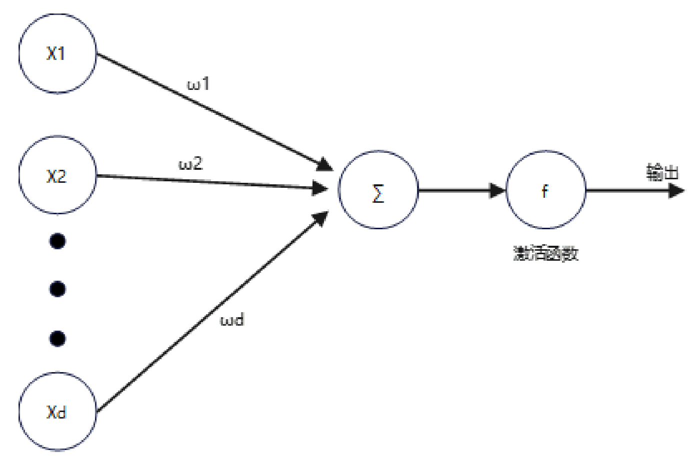


图1 神经元结构图

图1中包括了输入x_i，权重ω_i，节点阈值θ，激活函数$f(x)$以及输出y。神经元输出y可以用式(1)表示：

$$y = f\left(\sum_{i}^{d} \omega_i x_i - \theta\right) \tag{1}$$

对于神经元可以这么进行理解，在求和节点上对输入的信号进行加权求和，之后将加权求和之后的输入信号与节点阈值相减，最后通过激活函数将相减后的值传输给输出。

2.2 *激活函数*

激活函数有许多种，例如Sigmoid函数、Tanh函数、ReLU函数等。它们的功能主要是将输入的信息进行非线性映射传输到后一个节点，可以帮助神经网络学习输入数据中的复杂内容，增强神经网络的学习能力。选取合适的激活函数会帮助神经网络更好的学习，得到更高的训练效率以及正确率，同理，不合适的激活函数也会让神经网络的效率及正确率下降，有时会陷入到局部最优解。

下面将对几种常见的激活函数进行介绍。

Sigmoid函数的公式如式(2)所示：

$$f(z) = \frac{1}{1 + e^{-z}} \tag{2}$$

函数的图像如图2所示：

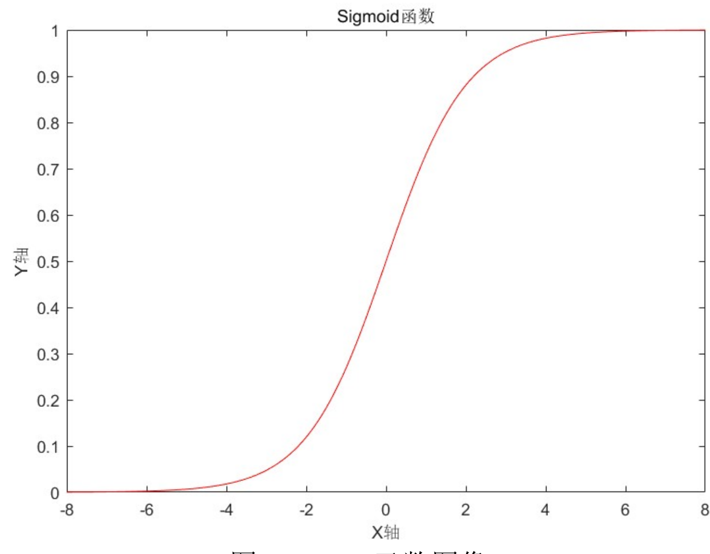


图2 sigmoid函数图像

从图像中可以看出，Sigmoid函数可以把输入的值变换为值域为[0,1]之间的输出，当输入的值十分小时输出为0，输入的值十分大时输出为1，这也导致了使用Sigmoid函数的时候，输入在接近这一区域的时候会饱和，也就导致了梯度的消失同时也会导致神经网络几乎不进行学习。另一点，在使用Sigmoid函数的时候，它的输出不是零中心的，会导致训练模型的收敛速度变慢。

Tanh函数是基于上述的Sigmoid的缺点进行改进的，其公式如式(3)所示：

$$\tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \tag{3}$$

它的函数图像如图3-3所示：

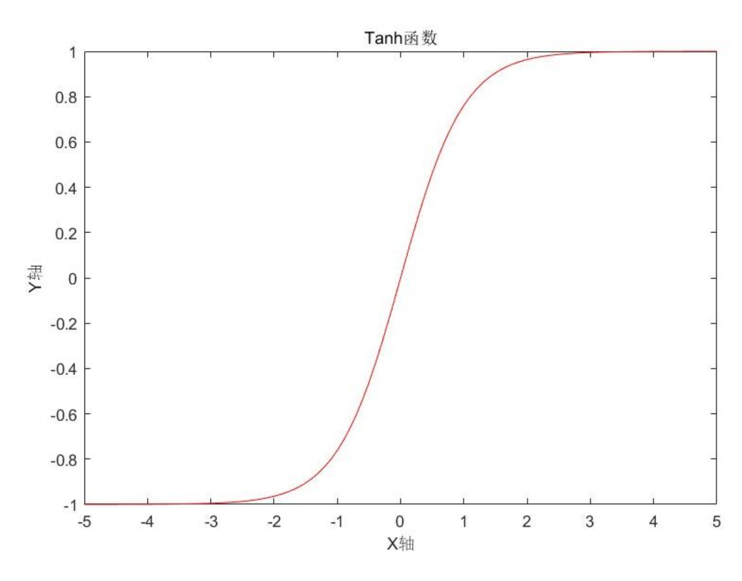


图3 Tanh函数图像

Tanh函数与Sigmoid函数一样，存在饱和的问题，不过不同的是，Tanh将输入的值变换为值域为[-1,1]之间进行输出，这也说明了Tanh函数的输出是零中心的，在运用过程中，Tanh函数的性能也会比Sigmoid函数表现更好。

ReLU函数的公式如式(4)所示：

$$\sigma(x) = \begin{cases} \max(0,x) & ,x \geq 0 \\ 0 & ,x < 0 \end{cases} \tag{4}$$

ReLU图像如图4所示,。

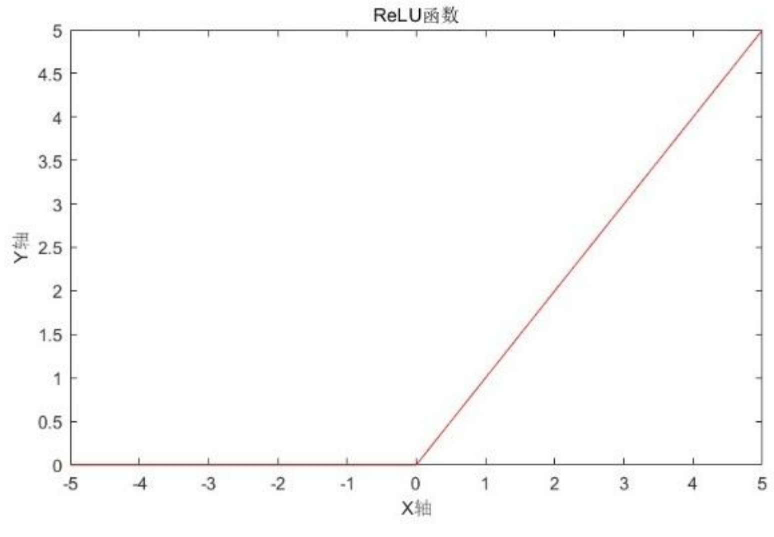


图4 ReLU函数的图像

从图中可以看出，ReLU函数在输入为负的时候，函数的输出也为0，这也导致了在神经网络进行传输的时候，有部分的神经元有可能不会被激活，也就起不到更新参数的作用，同样的ReLU函数和Sigmoid函数一样也不是零中心的。

在不同情况下需要选择不同的激活函数，激活函数好或坏，不能凭感觉定论，需要根据不同问题的不同性质，为神经网络选择更合适的激活函数。

2.3 前馈神经网络

前馈神经网络主要包括三个部分，输入层、输出层以及隐藏层，分为单隐层前馈神经网络以及多隐层前馈神经网络，二者区别在于隐藏层的层数[4]。

图5为单层前馈神经网络的结构。

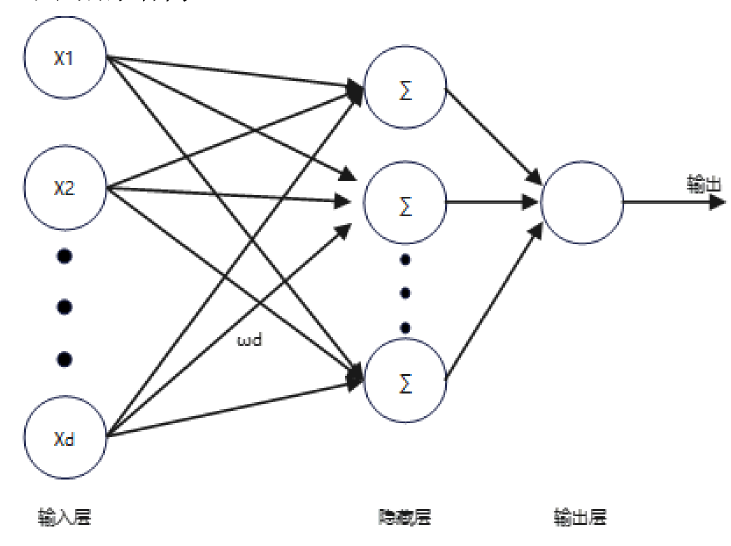


图5 单层前馈神经网络的结构

一个前馈神经网络可以包含三种节点：

1）输入节点（Input Nodes）：也称为输入层，输入节点从外部世界提供信息，。在输入节点中，不进行任何的计算，仅向隐藏节点传递信息。

2）隐藏节点（Hidden Nodes）：隐藏节点和外部世界没有直接联系。这些节点进行计算，并将信息从输入节点传递到输出节点。隐藏节点也称为隐藏层。尽管一个前馈神经网络只有一个输入层和一个输出层，但网络里可以没有也可以有多个隐藏层。

3）输出节点（Output Nodes）：输出节点也称为输出层，负责计算，并从网络向外部世界传递信息。

在前馈网络中，信息只单向移动——从输入层开始前向移动，然后通过隐藏层，再到输出层。在网络中没有循环或回路。

前馈神经网络中的不同层的神经元节点无法互联，每个神经元节点只负责接收处理前一层传输的信息，以及传输信息到下一层，前馈神经网络为输入层到输出层的单向传播。

2.4 BP神经网络

BP神经网络主要有两部分组成，一是正向传播，二是反向修正。在正向传播过程中，样本从输入层传播到隐藏层，在隐藏层经激活函数激活后再传输到输出层[5]。反向修正过程则是比较输出层的输出与期望输出，若不符合期望输出，则将实际值与期望值的差作为调整的信号，一层一层不断地进行反向传播，对连接权值进行调整，经过不断地正向以及反向传播，优化连接权值，当神经网络输出与期望输出的差值达到预期或者达到迭代次数之后，便完成了BP神经网络的训练。BP神经网络的流程图如图6所示：

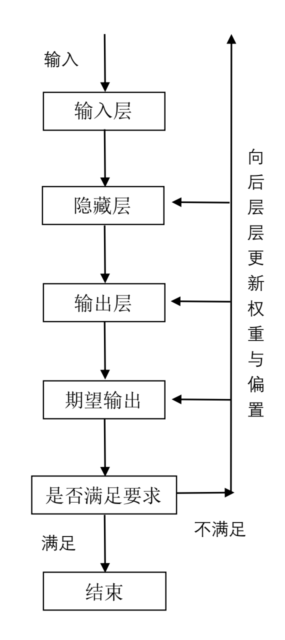


图6 BP神经网络的流程图

BP网络能学习和存贮大量的输入-输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程。它的学习规则是使用最速下降法，通过反向传播来不断调整网络的权值和阈值，使网络的误差平方和最小。其主要的特点是：信号是正向传播的，而误差是反向传播的。

但是BP神经网络有时会陷入局部极值，同时其算法的收敛速度较慢，需不断地进行迭代才能够达到良好的训练效果，所以，BP神经网络仍有一定的局限性[5]。

2.5 多层感知机

MLP是一种非线性分类器，是由输入层、隐含层(一层或多层)以及输出层构成的神经网络模型。图7为多层感知机的网络结构。

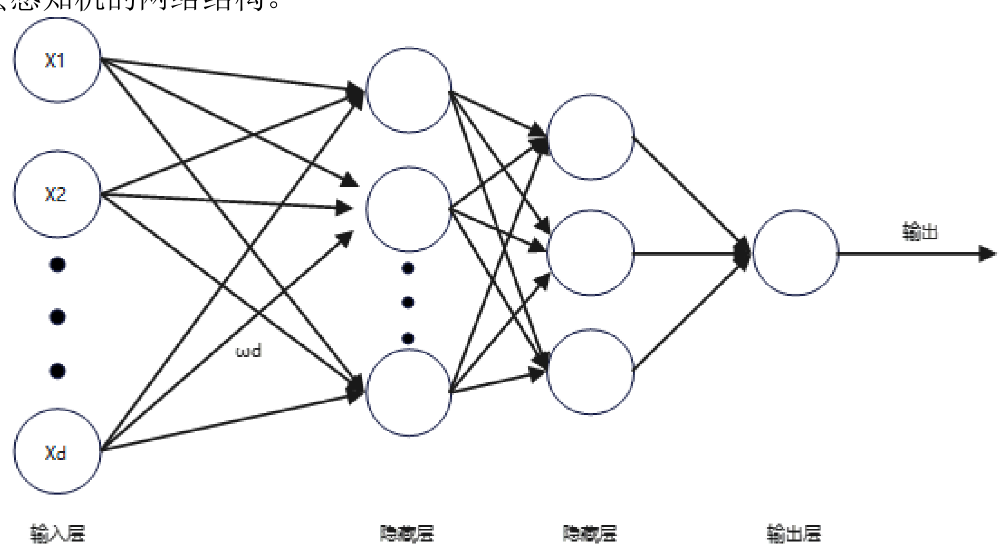


图7 多层感知机的网络结构

输入层神经元接收输入信号，层与层之间是全连接，每个连接都有一个连接权值，同层间的神经元互不相连。

MLP是一种多层神经网络，它的训练过程是这样的：首先，将一个特征向量作为输入，然后依次经过每一层的隐含层，每一层都会根据权重和激活函数计算出一个输出，并将其作为下一层的输入，直到最后一层的输出层。

通常，MLP使用BP反向传播算法来训练，这个算法的思想是通过错误来学习。当人工神经网络的输出与真实值不一致时，监督者会对网络进行调整。MLP由多层节点组成，包括输入层，中间隐藏层和输出层。相邻层节点之间有权重连接。学习的目标是为这些连接找到合适的权重。这些权重会影响输入向量到输出向量的映射。在监督学习中，训练集是有标签的。这意味着对于每个输入，我们都知道应该得到什么输出。MLP训练过程大致如下：

（1）所有边的权重随机分配；

（2）前向传播：利用训练集中所有样本的输入特征，作为输入层，对于所有训练数据集中的输入，人工神经网络都被激活，然后经过前向传播，得到输出值。

（3）反向传播：利用输出值和样本值计算总误差，再利用反向传播来更新权重。

（4）重复（2）~（3），直到输出误差低于制定的标准。

MLP是一种具有前向结构的ANN人工神经网络，它可以处理非线性可分离的问题，是一个值得深入研究的模型。MLP的实现需要用到BP反向传播算法，这是一种通过调整网络权重来最小化误差的算法。MLP可以使用任何形式的激活函数，但为了让BP算法有效地学习，激活函数必须是可微的。

MLP有很多的优点，如高度的并行处理，高度的非线性全局作用，良好的容错性，具有联想记忆功能，非常强的自适应、自学习能力等。但也有一些缺点，如学习速度慢，容易陷入局部极值，学习可能会不够充分等。MLP有很广泛的应用范围，扩展性也强，可以用于语音识别、图像识别、机器翻译等领域。

3 深度学习

深度学习(DL)是机器学习(ML)的一个新分支，它让机器学习更接近它的原始目标——人工智能(AI)。深度学习是从样本数据中挖掘内部规律和层次结构，这些信息有助于理解文字、图像和声音等数据。它的最终目的是使机器具备人类的分析和学习能力，能够识别文字、图像和声音等数据。深度学习是一种复杂的机器学习方法，在语音和图像识别方面的表现，远超过之前的相关技术。

3.1 卷积神经网络

卷积神经网络（Convolutional Neural Networks）是一种深度学习模型或类似于人工神经网络的多层感知器，常用来分析视觉图像。

一个卷积神经网络主要由以下5层组成：数据输入层，卷积计算层，ReLU激励层，池化层，全连接层，卷积神经网络的结构图如图8所示。

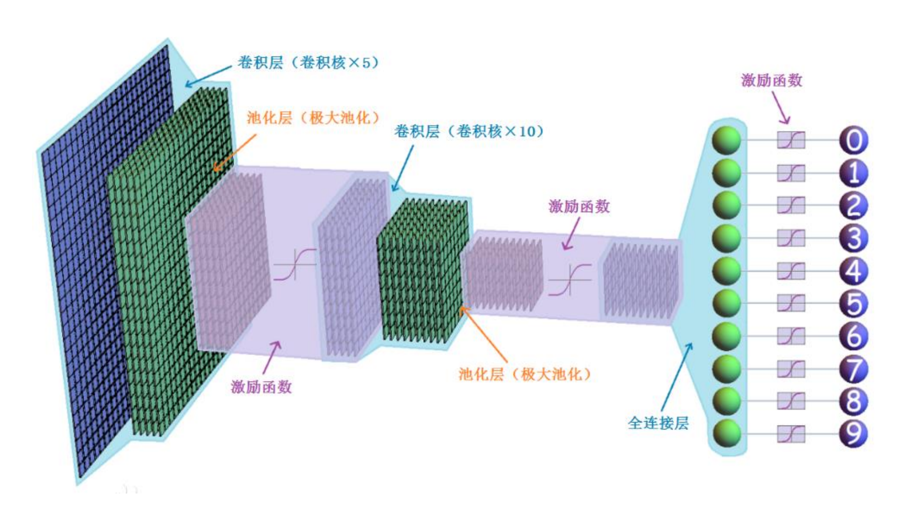


图8 卷积神经网络结构图

卷积层和池化层一般会取若干个，采用卷积层和池化层交替设置，即一个卷积层连接一个池化层,池化层后再连接一个卷积层依此类推.由于卷积层中输出特征面的每个神经元与其输入进行局部连接，并通过对应的连接权值与局部输入进行加权求和再加上偏置值,得到该神经元输入值，该过程等同于卷积过程，CNN也由此而得名[6]。

它们的作用可以简单理解为：卷积层负责提取图像中的局部特征；池化层用来大幅降低参数量级(降维)；全连接层类似传统神经网络的部分，用来输出想要的结果。

1. 卷积层

卷积层的运算过程是用一个卷积核扫完整张图片，可以理解为使用一个过滤器（卷积核）来过滤图像的各个小区域，从而得到这些小区域的特征值。

在CNN 结构中，深度越深、特征面数目越多，则网络能够表示的特征空间也就越大、网络学习能力也越强，然而也会使网络的计算更复杂，极易出现过拟合的现象。因而，在实际应用中应适当选取网络深度、特征面数目、卷积核的大小及卷积时滑动的步长，以使在训练能够获得一个好的模型的同时还能减少训练时间。

总的来说，卷积层的通过卷积核的过滤提取出图片中局部的特征，跟上面提到的人类视觉的特征提取类似。

2. 池化层

池化层紧跟在卷积层之后，同样由多个特征面组成，它的每一个特征面唯一对应于其上一层的一个特征面，不会改变特征面的个数，简单说就是下采样，可以大大降低数据的维度。

对卷积后的特征图用选定的步长依次对一定大小的图像区域进行降采样操作，取出该区域的典型特征值，特征值的选取通常有三种：最大值、平均值和随机值［２３］。对应的池化方法分别为最大化池化、平均池化和随机池化，随机池化方法具有随机性，按照概率矩阵依大小随机选择。池化操作将池化区域内的多个值变为一个值，特征图的宽度和高度均降为原来的几分之一，减少了卷积层输出的特征向量，极大地减少了参数数量和卷积网络中的计算量，同时能保证数据的平移不变性，池化层相比卷积层可以更有效的降低数据维度，这么做不但可以大大减少运算量，还可以有效的避免过拟合。

3. 全连接层

在CNN结构中，经多个卷积层和池化层后，连接着１个或１个以上的全连接层．与MLP类似，全连接层中的每个神经元与其前一层的所有神经元进行全连接，将输入层图像特征经多层特征提取再经高度融合后，得到图像的高层表示，最后经回归分类模型计算得到相应的响应值送到输出层。

CNN 很擅长处理图像，可以用在图像分类、检索，目标定位检测，目标分割，人脸识别，骨骼识别等领域。卷积神经网终在有监督的学习中已有广泛的应用，但有监督的学习需要大量的图像标注,耗费工作量大，如何在这此工作上提升效率还有待研究。对于原始图像中含有一定噪声的图片,进一步提高模型的抗噪能力及泛化能力，也是待解决的问题。对于快速性要求较高的任务及提高图像识别的速度场景需要增加卷积神经网终的实时性能，设计出满足实时性高的轻量级卷积神经网络，同时兼顾模型准确性能。

3.2 自动编码器

自动编码器是一种利用神经网络来学习输入数据的隐含特征的无监督式学习模型。它的训练目标是使得神经网络的输出能够尽可能地接近输入，而不需要任何额外的标签信息。它由两个主要的部分组成：编码器和解码器，编码器的作用是将输入数据X压缩成一个低维的编码表示X^E，解码器的作用是将编码表示X^E还原成一个重构输出X^R，其结构如图9所示。

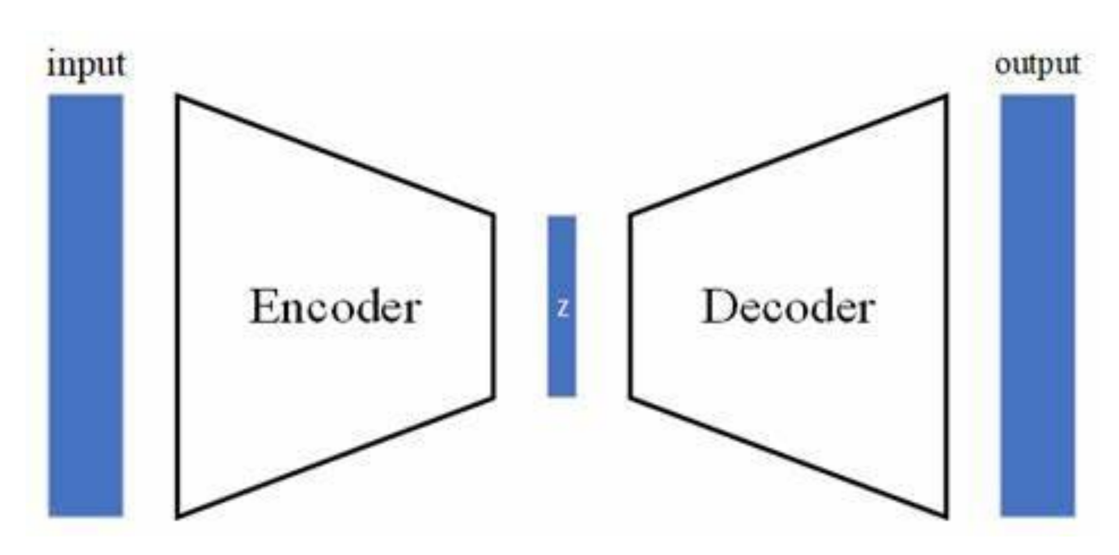


图9 自编码器的结构图

自动编码器能够从输入数据中提取隐含的特征，这个过程叫做编码，然后利用这些特征来重建输入数据，这个过程叫做解码。直观地说，自动编码器可以实现特征降维的功能。除此之外，自动编码器得到的新特征还可以作为有监督学习模型的输入，因此自动编码器也可以用来做特征提取。

自动编码器有三大特点：

1、自动编码器是数据相关的，这意味着自动编码器只能压缩那些与训练数据类似的数据。比如，使用人脸训练出来的自动编码器在压缩别的图片，比如树木时性能很差，因为它学习到的特征是与人脸相关的。

2、自动编码器是有损的，意思是解压缩的输出与原来的输入相比是退化的，MP3，JPEG等压缩算法也是如此。这与无损压缩算法不同。

3、自动编码器是从数据样本中自动学习的，这意味着很容易对指定类的输入训练出一种特定的编码器，而不需要完成任何新工作。

搭建一个自动编码器需要完成下面三样工作：搭建编码器，搭建解码器，设定一个损失函数，用以衡量由于压缩而损失掉的信息。编码器和解码器一般都是参数化的方程，并关于损失函数可导，典型情况是使用神经网络。编码器和解码器的参数可以通过最小化损失函数而优化，例如SGD。

它的优点是泛化性强，无监督不需要数据标注；缺点是针对异常识别场景，训练数据需要为正常数据。

3.4 变分自编码器

变分自编码器（VAE）是一种深度生成模型，由 Kingma 等人于 2014 年提出，基于变分贝叶斯（VB）推断的方法来构建生成式网络。它不同于传统的自编码器，它不是用数值来表示潜在空间，而是用概率来描述潜在空间的状态，在生成数据方面有很大的优势[7]。

图10为变分自编码器的结构图。

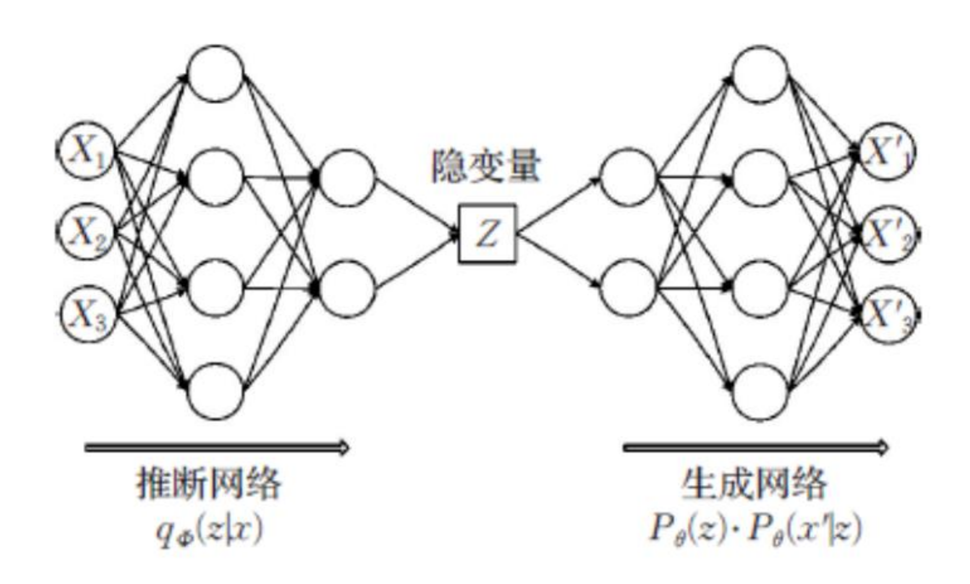


图10 变分自编码器的结构图

变分自编码器与自动编码器由编码器与解码器两部分构成相似，VAE利用两个神经网络建立两个概率密度分布模型：一个用于原始输入数据的变分推断，生成隐变量的变分概率分布，称为推断网络；另一个根据生成的隐变量变分概率分布，还原生成原始数据的近似概率分布，称为生成网络。

但是，为了引入隐空间的某些正则化，对编码-解码过程进行了一些修改：不是将输入编码为隐空间中的单个点，而是将其编码为隐空间中的概率分布。然后对模型进行如下训练：首先，将输入编码为在隐空间上的分布；二，从该分布中采样隐空间中的一个点；第三，对采样点进行解码并计算出重建误差；最后，重建误差通过网络反向传播。

由于VAE整体结构与自编码器 AE结构类似，因此被称为变分自编码器。但是，VAE 的作用原理和AE的作用原理完全不同，VAE的“编码器”和“解码器”的输出都是受参数约束变量的概率密度分布，而不是某种特定的编码。

4 结论

神经网络是一种受人脑启发的人工智能模型，它由许多神经元构成，可以从数据中学习复杂的函数和模式。神经网络有多种形式和用途，其中卷积神经网络（CNN）是一种专门处理图像和空间数据的神经网络，它利用局部连接和权值共享的方法，减少参数数量，提取图像的特征和语义。CNN的优点是可以有效地处理高维数据，提高图像识别、检测、分割等任务的性能；缺点是需要大量的训练数据和计算资源。自编码器（AE）是一种无监督的神经网络，它的目的是通过一个低维的隐层，重建输入数据，从而学习数据的有效表示。AE的优点是可以在没有标签的情况下，发现数据的潜在结构，进行特征提取、降维、去噪等任务；缺点是可能出现信息丢失、过拟合、无法生成新的数据样本等问题。变分自编码器（VAE）是一种基于概率的自编码器，它将输入数据映射到一个潜在的随机变量，从而可以生成新的数据样本。VAE的优点是可以实现数据的生成、插值、变换等任务，具有更强的泛化能力；缺点是需要引入额外的随机噪声，可能导致重构误差增大，生成的数据质量不高。神经网络、卷积神经网络、自编码器和变分自编码器都是深度学习的重要工具，它们在图像处理、自然语言处理、计算机视觉等领域有着广泛的应用和研究价值。

参考文献：

[1] 张松兰. 基于卷积神经网络的图像识别综述 [J]. 西安航空学院学报, 2023, 41 (01): 74-81. DOI:10.20096/j.xhxb.1008-9233.2023.01.013

[2] Nguyen VanCuong, Hoang DuyTang, Tran XuanToa, et al. A Bearing Fault Diagnosis Method Using Multi-Branch Deep Neural Network[J]. Machines, 2021, 9(12):345-345

[3] Afrasiabi Shahabodin, Mohammadi Mohammad, Afrasiabi Mousa, et al. Modulated Gabor filter based deep convolutional network for electrical motor bearing fault classification and diagnosis[J]. IET Science, Measurement & Technology, 2021, 15(2):154-162

[4] 李娜,刘冰,王伟.基于单隐层前馈神经网络的优化算法[J].科学技术与工程,2019,19(01):136-141

[5] 邓长城,李辉,苏金虎,马晓杰. 基于 BP 神经网络的轴承温度预测[J]. 河南工学院学报,2022,30(03):1-6

[6] 周飞燕,金林鹏,董军. 卷积神经网络研究综述 [J]. 计算机学报, 2017, 40 (06): 1229-1251.

[7] 翟正利,梁振明,周炜等. 变分自编码器模型综述 [J]. 计算机工程与应用, 2019, 55 (03): 1-9.